

Lights, Camera, Interaction

A virtual cinematography system relinquishes lighting and camera control to a digital director

DIANA PHILLIPS MAHONEY

Without its unique cinematography, *The Blair Witch Project* would look like a nature documentary, says MIT Media Lab researcher Bill Tomlinson in stressing the importance of "expressive" camera and lighting techniques to the overall look and feel of a film. "In a movie, you can tell how the characters are feeling by how each shot is composed and how the shots are strung together. Camera and lighting work can change the whole tone of a film." The gut-churning *Blair Witch Project*, for example, relies on an endless stream of violent, jerky camera movements and ominous lighting to build a sense of horror in the audience and to imply terror that the content itself doesn't explicitly reveal.

The ability to generate such emotion through cinematography is an idea that Tomlinson and his colleagues in the Media Lab's Synthetic Characters Group, headed by professor Bruce Blumberg, want to bring to interactive virtual worlds. Toward this end, the researchers have developed a behavior-based system that interacts with the emotions and motivations of autonomous characters in a scene and automatically chooses optimal shots and lighting scenarios to best represent the mood of the environment.

The expressive-cinematography technology is an outgrowth of the Synthetic Character Group's work in autonomous character design, which is modeled on principles of animal behavior, says

Tomlinson. "Animals are really good at making sense of complicated and changeable environments. They have likes and dislikes, needs and fears, that help them figure out what actions they should take."

The digital autonomous characters mimic this state in the virtual world by following a set of simple rules encompassing motivations, emotions, and action-selection mechanisms that parallel those of animal behavior. The rules tell the characters how to behave in a given environment. "Just as a wolf might be hungry for a hunk of meat, our autonomous characters are hungry for whatever food resources are available in their world," says Tomlinson. "Such motivations are combined with emotional states, which influence the action-selection mechanism: Should I chase that yummy-looking chicken, who might squawk at me and scare me, or should I try to poach her eggs instead?"

The virtual cinematography system further extends the autonomous-character paradigm, by relying on a "CameraCreature" who operates behind the scenes to direct the virtual camera views and lighting placement. The CameraCreature is built on the same architecture as the autonomous characters that populate the virtual environments. "It works the same way as the rest of our characters. It just has a different definition of what tastes good. Rather than having a taste for virtual ham and eggs, it's hungry for well-framed shots."

To satisfy its appetite, the CameraCreature automatically manipulates such elements as camera position and motion, timing of edits, transitions between shots, and various lighting effects. As new characters emerge and novel scenarios unfold, the camera view and lighting change dynamically to best reflect the story development. Because it shares the same underlying rules-based structure as the actors, the CameraCreature is able to communicate with the other characters. "It can query such elements from each character as emotional state and position. In addition, characters are able to make requests of the system if they think they're doing something that should be on camera," says Tomlinson.

The CameraCreature's behavior system consists of four elements: sensors, emotions, motivations, and actions. The sensors extract pertinent information

The antics of K.F. Chicken, the star of an interactive experience called *Swamped!* are filmed by an autonomous CameraCreature, whose decisions are driven by the emotional elements of a given scene.





Happy synthetic characters make for a happy CameraCreature, which in turn makes for bouncy, sweeping camera motions and bright lighting. Should the mood of the digital characters change, the camera and lighting views would be modified to best reflect the new mood.

from characters in each scene, including position, orientation, size, motivation, and emotion. These factors become part of the system's knowledge base for making lighting and view decisions. Similarly, the emotions of the characters affect those of the CameraCreature, which influence the shot decision. The impact of a character's mood on that of the CameraCreature depends on a number of predetermined factors, including the "sensed" importance of the character, how much screen time it's had, and how strong the perceived emotion is.

The CameraCreature's own emotion structure consists of six primary emotions: happy, sad, angry, surprised, fearful, and disgusted. Each of these elements is associated with specific base levels, inputs, gains, and rates of change to determine the CameraCreature's emotion at every time step, which largely defines its "professional" behavior. For example, when it's happy, it might want bouncy, sweeping camera motions and bright lighting. Conversely, a sad CameraCreature might prefer lugubrious motion and dimmed lights.

In addition to evaluating emotions, the cinematography system also considers motivations for the type of shot to use for a particular event. For example, based on the relative values of the inputs it's receiving from

the scene (through the characters' emotions and/or requests), the CameraCreature may determine an establishing shot is necessary to familiarize the participant with the environment, or that a close-up is warranted to focus on a specific, "important" character.

Ready, Set, Action!

Finally, the action-selection element is what motivates the CameraCreature's decisions. It is based on a hierarchical organization through which the "value" of a given action is weighed against that of a conflicting action. For example, at every time step, a character is faced with action choices. Options might include getting food, running from danger, or exploring the environment. Each action is defined by a routine, which sends a message to the code for calculating where the camera should be positioned and how it should be oriented. Because only one of the three actions can be accomplished at a time, the



actions battle it out based on their output values (motivations and emotions).

Under each general action category are a number of more specific actions. For example, under "get food" might be "go to the grocery store," "order a pizza," "have some cereal," and so forth. During the action-selection process, says Tomlinson, "the most general groups compete with each other to determine what basic category of action will be taken. Then the sub-groups of the winning action compete to find out which will win. This continues down the levels of the hierarchy until only one specific action remains, which the creature then performs." The values of all of the relevant behaviors are calculated at every time step.

Tomlinson presents an imaginary real-world scenario to illustrate the process. "Imagine we have a guy who's really excited to watch *Buffy the Vampire Slayer*, but he's hungry too. He has sensory information from the TV guide that *Buffy* just started. He also has a strong motivation to acquire food. It makes him unhappy—changes his emotional state—that he can't do both, but it is more important to him that he watch the opening credits. After the *Buffy* stimulus dies down—when the show ends—his hunger will probably be most important in his action-selection process, and he'll take off for Taco Bell."

The CameraCreature would process a similar chain of decisions to film the same scene, Tomlinson says. "It would have sensory information about where the guy and the TV are in the apartment. It would feel a motivation to show relationships between the guy, the TV, and the Taco Bell seen out the window by framing shots around them. It would have an emotional state, in that it would want to make the user feel the tension in the scene. And it would be prepared to change the decisions made by its action-selection mechanism depending on whether the guy decides to sit or go to Taco Bell."

Once the action-selection process has transpired, the CameraCreature acts on it, deciding where to put the camera, the direction in which to face it, and how it should move through the world. The decisions generally come down to the question of which actors to watch. "The camera doesn't necessarily have to cover everything that happens, since in a complex world it's impossible to show everything, but it does need to show everything that's important. The trick is figuring out what's important,"



says Tomlinson. Because "important things" are usually those most closely related to the human participant, the CameraCreature's shot decision is usually weighted in favor of the actor with whom the participant is interacting. "If a shot is really cool but doesn't matter to the interaction the user is engaged in, maybe the system shouldn't show it."

However, the camera does need to be ready to cut to another character if it's performing an interesting, relevant action, says Tomlinson. "Since the characters are free to take a wide variety of actions at any time, the camera has to be prepared to deal with these possibilities. It doesn't have the opportunity to say, 'Wait, that shot was lame. Can I try it again?'"

In addition to camera location, the CameraCreature must also decide on a camera angle with respect to the coordinate system of the target actor or actors. Different angles serve different purposes and are selected based on the emotion, motivation, and action-selection criteria. Wide angles orient the participant to the virtual world. Close ups are useful for

expressive shots to highlight specific emotions. Navigational angles let the camera track a character's motion. Each of the angles has parameters that let it adjust camera and target positions.

The camera's movement through the virtual world is achieved through a parametrical system of virtual springs and dampers. The CameraCreature manipulates the parameters of the spring dynamics system to enable effects that are tied to its emotions. An angry CameraCreature might cause quick, sharp movements, a sad one might move in slow arcs.

As it does with the camera work, the CameraCreature automatically evaluates the emotional, motivational, and action-selection elements of a scene to determine optimal lighting conditions, then configures the global and personal lights accordingly.

Currently, the autonomous cinematography system is running entirely on PCs. Most of the programming is done in Java, although some C and C++ code is implemented as well. The system is one piece of the larger intelligent-systems puzzle the Media Lab researchers are putting together,

the goal of which is to create and display characters that make participants feel as if they're interacting with living characters.

Among the Synthetic Character Group's current research pursuits are advanced interactive lighting design, enhanced adaptive story development, and multi-user interaction. Longer-term objectives include allowing the camera to cheat reality—to halt a course of events for later coverage, for example, or to move characters and scene elements for better framing. Also, the researchers hope eventually to develop a smart camera that is capable of learning such things as the nature of characters it has worked with before and the types of shots that have worked well.

The goal of this research, stresses Tomlinson, is to explore technologies that might be five or ten years ahead of industry, as a way to plant seeds for future development. "Our job is to make stuff that might change the way our sponsors think about their industry."

Diana Phillips Mahoney is chief technology editor of Computer Graphics World.

On the Right Track

A unique optical tracking system gives users greater freedom to explore virtual worlds

The success of any virtual-reality experience depends on the user's sense of presence in the virtual environment. While believable graphics and familiar modes of interaction are important elements to achieving such presence, both are of little value if the user's navigation through the digital scene feels slow, stilted, or somehow out of sync, or if the user feels prohibitively constrained in space—all effects of less-than-perfect motion tracking.

A motion-tracking system is the hardware/software configuration through which a user's head position and orientation in a scene are communicated to the computer, which in turn adjusts the displayed image to reflect changes in the user's location. This is generally achieved through one of five types of motion-tracking technologies: mechanical, electromagnetic, inertial, optical, or acoustic.

An ideal tracking system is one that is able to generate position and orientation computations in real time without compromising the graphical display or the user's freedom of motion. Unfortunately, such an ideal has yet to be realized. In reality, most tracking systems are forced to make technology tradeoffs.

With mechanical tracking, for instance, the tracking hardware is physically attached to the object or person being tracked, so while such devices offer high accuracy, they restrict movement. Electromagnetic tracking measures the strength of the



The HiBall technology is packaged in a golf-ball sized device fitted with specialized optical sensors and lenses that can be head-mounted to track the user's view perspective through a virtual or augmented reality scene. Housed inside the device are stacked, folded, flexible PC boards.